Housekeeping Rules

Attendees will be entered into the webcast on mute and will remain muted at all times.

Use the Chat Box on the webcast or the live Q/A feature on the conference app to share a question or comment and to communicate with presenters and moderators during the Summit.

If you have technical issues, please seek assistance from your local technical support.
Methodologic Issues for PPI Studies
Considerations for Choosing the PPI Method for Medical Device Evaluation

David Gebben PhD
Calvin University
Calvin email: dig5@calvin.edu
Disclaimer

This presentation represents the personal opinions of the speaker and does not represent the views or policies of FDA.

No conflicts of interest to declare.

I am employed by Calvin University.
Considerations for Choosing & Using a PPI Method

• The selected method should support the objective of the study, such as to inform:
  – Endpoint selection
  – Benefit-risk decisions (premarket or postmarket)

• Consider the total product lifecycle (TPLC) since different methods may be better suited at different points along the TPLC
  – If endpoint selection is the objective, then the PPI study would need to be completed prior to the clinical trial

• Robust analysis of results
Research Design Process

Figure 1. Summary of qualitative steps for preference survey development

Step 1: Identify relevant research question
- Potential pre-submission meeting with FDA

Step 2: Define study result of interest
- Potential pre-submission meeting with FDA

Step 3: Define preference elicitation method and study design

Step 4: Identify attributes and attribute levels
- Potential pre-submission meeting with FDA

Step 5: Develop the preference survey

Step 6: Pretest preference survey
- Potential pre-submission meeting with FDA

The Research Question—Where it All Starts

A clear research question is:
- Well-defined
- Narrow to prevent scope creep
- Aligned with the study’s objective
- Specified clearly in submission
PPI Elicitation Methods and Study Designs

Preference elicitation methods

- Ranking: order of preference
- Rating: magnitude of preference
- Threshold technique: individual trade-offs
- Discrete choice: multiple trade-offs

Discrete Choice Experiments (DCE)

- Currently, this is the most familiar method
  - Allows evaluation of multiple attributes at once
  - Can inform endpoint selection prior to clinical trials
  - Can inform benefit-risk analysis
Another familiar method
• Can inform endpoint selection for clinical trials
• Allows for a benefit-risk analysis
• Can only evaluate one attribute at a time
• Potentially less burdensome for the respondent than DCE which may be more appropriate for certain populations (e.g. pediatric)
Best Worst Scaling (BWS)

• Could be used to inform the prioritization of endpoint selection
• Could also be used to inform prioritization of device features in development stage
Other Potential Methods

• Swing Weighting
  – Possible use with rare or hard to reach populations

• Focus groups
  – Could be used with rare or hard to reach populations
  – When quantitative research would not be possible or needed based on the research question, patient population, or stage of the research
## Attribute Considerations

<table>
<thead>
<tr>
<th>Framing Effects</th>
<th>• Asking “Which would you choose” is not equal to “Which is better”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to Decision</td>
<td>• Decision proximity impacts prevention versus promotion focus</td>
</tr>
</tbody>
</table>
| Number of Attributes | • Has to be relevant  
                      | • Consider the design space  
                      | • Too many choices creates more opportunities for deferring the decision |
| Number and Spacing of levels | • Reflect the clinical range  
                         | • Spaced so that patients do not inadvertently recode those levels  
                         | • Ensure levels do not overlap statistically |
Robust Analysis of Results

- The analysis of results should be robust
  - Determine sources of uncertainty
  - Sensitivity of analysis
- Heterogeneity Considerations
Robust analysis

• The FDA PPI Guidance does not identify a specific analysis method

• Sound and robust analyses should answer the following questions:
  – Did the analysis identify the sources of uncertainty?
  – Did you consider subpopulations that could be accounting for increased variability?
  – Is the analysis robust to using different modeling assumptions?

• Presentation of the analysis should be clearly identified and explained in the submission
Heterogeneity Considerations—The Struggle Is Real

• Patient characteristics that could be associated with heterogeneity include
  – Time with condition
  – Severity of condition
  – Proximity to choice
  – Socio-demographic and cultural factors

• Consider it during the design phase as well as analysis stage

• May not be able to completely address it - but should acknowledge it

Do not skip the first step in the research – it is an opportunity to get FDA feedback.

We have more than just a hammer in the toolkit so do not think of every question as a nail.

Whatever analysis is chosen, it should be:
- Robust
- Addressing the research question
- Relevant to the medical device regulatory decision
BRAVE Initiative

*Benefit - Risk Assessment Valuation & Evidence*

Understanding Patient and Caregiver Treatment Preferences

Parent Project Muscular Dystrophy

JOIN THE FIGHT. END DUCHENNE.

Ryan Fischer
SVP Community Engagement
The patient voice shaping research and policy

Advocacy and Legislation
- Advocating for PFDD and shaping rare disease policy and legislation

Research and Reports
- Interpreting, publishing, disseminating data
- Community facing reports

Guidance
- First patient group to submit a community led draft guidance to FDA
Community-Engaged Approach – Key Principles

Advocacy-led initiative

Stakeholder-engaged process

Data owned by the advocacy community

Dissemination through patient group
## Testing different methods on our population (capacity building)

<table>
<thead>
<tr>
<th>Study</th>
<th>Topic</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pilot (2013)</td>
<td>Treatment preferences</td>
<td>Best-worst scaling case 2</td>
</tr>
<tr>
<td></td>
<td>Treatment preferences</td>
<td>Conjoint analysis validation</td>
</tr>
<tr>
<td></td>
<td>Parental worry</td>
<td>Best-worst scaling case 1</td>
</tr>
<tr>
<td></td>
<td>Pulmonary outcomes</td>
<td>Conjoint analysis validation</td>
</tr>
<tr>
<td></td>
<td>Symptom treatment priorities</td>
<td>Best-worst scaling case 1</td>
</tr>
<tr>
<td>3. Advancing methods   (2016)</td>
<td>Treatment preferences + uncertainty</td>
<td>Discrete-choice experiment</td>
</tr>
<tr>
<td>4. Gene therapy (2018)</td>
<td>Clinical trial decision making priorities</td>
<td>Qualitative and Best-worst</td>
</tr>
<tr>
<td></td>
<td>Risk of Death</td>
<td>Threshold</td>
</tr>
<tr>
<td>5. Global Study (2019)</td>
<td>Treatment preferences + uncertainty</td>
<td>Discrete-choice experiment</td>
</tr>
</tbody>
</table>

### Collaborators
- Dr. John Bridges (The Ohio State)
- Nonie Crossnohere (JHSPH)
- Dr. Holly Peay (RTI International)
- Pfizer
- Everylife Foundation
- Solid Biosciences
- Santhera Pharmaceuticals

### Funding
Duchenne Muscular Dystrophy

- X-linked, pediatric onset neuromuscular disease
- Lack of critical protein called dystrophin
- Incidence: 1:4600 boys (30% spontaneous)
- Approximately 20,000 people in US
- Diagnosis: 3-5 years of age
- Predictable course of progressive loss of function
- 100% lethal - most die by mid twenties
- 4 approved therapies, very active pipeline

- One of the largest genes in the human genome (2.4m base pairs/79 Exons)
- Caused by mutations in the dystrophin-encoding DMD gene
Pilot Study
Caregiver Preferences (2013)
BW Case 1
Experiment 1 - Worry Prioritization (BW Scaling 1)

Aim: Understand prioritization of parental worries related to Duchenne progression and if worry prioritization varies based on ambulation status

**Themes and Worries**

<table>
<thead>
<tr>
<th>Medical Concern</th>
<th>Child Affect/emotion</th>
<th>Parent Wellbeing</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Child Getting Weaker</td>
<td>My Child feeling happy</td>
<td>Managing my uncertainty about child's future</td>
</tr>
<tr>
<td>Getting the Right Care for my child</td>
<td>My child having good friends</td>
<td>Being a good enough parent for child</td>
</tr>
<tr>
<td>Affording Care for my child</td>
<td>My child not being able to express worries</td>
<td>Handing the emotional demands of Duchenne</td>
</tr>
<tr>
<td>My child missing out on new treatments</td>
<td>My child feeling like a burden to family</td>
<td>Having time for myself</td>
</tr>
</tbody>
</table>

**Family and social**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The wellbeing of my other children</td>
<td></td>
</tr>
<tr>
<td>My child becoming independent over time</td>
<td></td>
</tr>
<tr>
<td>Effect of Duchenne on my close relationships</td>
<td></td>
</tr>
<tr>
<td>Feeling isolated from other family</td>
<td></td>
</tr>
</tbody>
</table>

**Example task**

*In the past 7 days, choose which you have been most worried about and which you have been least worried about.*

<table>
<thead>
<tr>
<th>Most worried</th>
<th>Worries</th>
<th>Least worried</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>My child getting weaker</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>Being a good enough parent</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>Affording care within our budget</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>Having time for myself</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>Feeling isolated from other family</td>
<td>o</td>
</tr>
<tr>
<td>o</td>
<td>My child feeling happy</td>
<td>o</td>
</tr>
</tbody>
</table>
## Experiment 2 - Treatment Scenarios (BW Scaling 1 and 2)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effect on muscle function</strong></td>
<td>• Stops the progression of weakness</td>
</tr>
<tr>
<td></td>
<td>• Slows the progression of weakness</td>
</tr>
<tr>
<td></td>
<td>• Does not change progression of weakness</td>
</tr>
<tr>
<td><strong>Lifespan</strong></td>
<td>• 5 year gain in expected lifespan</td>
</tr>
<tr>
<td></td>
<td>• 2 year gain in expected lifespan</td>
</tr>
<tr>
<td></td>
<td>• No extra gain to expected lifespan</td>
</tr>
<tr>
<td><strong>Knowledge about the drug</strong></td>
<td>• 2 years of post-approval drug information</td>
</tr>
<tr>
<td></td>
<td>• 1 year of post-approval drug information</td>
</tr>
<tr>
<td></td>
<td>• No post-approval drug information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nausea</strong></td>
<td>• No increased chance of nausea</td>
</tr>
<tr>
<td></td>
<td>• Causes loss of appetite</td>
</tr>
<tr>
<td></td>
<td>• Causes loss of appetite with occasional vomiting</td>
</tr>
<tr>
<td><strong>Risk of bleeds</strong></td>
<td>• No increased risk of bleeds</td>
</tr>
<tr>
<td></td>
<td>• Increased risk of bleeding gums &amp; bruising</td>
</tr>
<tr>
<td></td>
<td>• Increased risk of hemorrhagic stroke and lifelong disability</td>
</tr>
<tr>
<td><strong>Risk of heart arrhythmia</strong></td>
<td>• No increased risk of heart arrhythmia</td>
</tr>
<tr>
<td></td>
<td>• Increased risk of harmless heart arrhythmia</td>
</tr>
<tr>
<td></td>
<td>• Increased risk of dangerous heart arrhythmia and sudden death</td>
</tr>
</tbody>
</table>

**Example task**

Choose the best thing and worst thing about this treatment. Then answer the question about whether you would use it for your child.

### Treatment

<table>
<thead>
<tr>
<th>Best</th>
<th>Treatment</th>
<th>Worst</th>
</tr>
</thead>
<tbody>
<tr>
<td>✰</td>
<td>Slows the progression of weakness</td>
<td>✰</td>
</tr>
<tr>
<td>✰</td>
<td>2 year gain in expected lifespan</td>
<td>✰</td>
</tr>
<tr>
<td>✰</td>
<td>1 year of post-approval drug information</td>
<td>✰</td>
</tr>
<tr>
<td>✰</td>
<td>Causes loss of appetite</td>
<td>✰</td>
</tr>
<tr>
<td>✰</td>
<td>Increased risk of bleeding gums and increased bruising</td>
<td>✰</td>
</tr>
<tr>
<td>✰</td>
<td>Increased risk of harmless heart arrhythmia</td>
<td>✰</td>
</tr>
</tbody>
</table>

---

**If this treatment were real, would you use it for your child?**

- ☐ Yes
- ☐ No
- ☐ I don’t know
Results - Caregiver Preferences for Emerging Treatments

Worries Experiment:
• Most worried about child getting weaker and getting good care for their child
• Ambulation status showed some differences but not highest prioritized items

Treatment Scenario:
• Caregivers prioritized protection of muscle function over all attributes including longer lifespan
• Slowing disease progression was a highly valued outcome
• Caregivers chose quality of life over quantity

Methods:
- Chose simplest method for transparency and ease of experiment
- “Bottom up” approach to attribute selection
- Understandable, acceptable dissemination to community and FDA were a primary objectives
Symptom Treatment Priorities & Meaningful Benefit in Pulmonary Function (2015)

BW Case 1 and 2
Patients and Caregivers
Experiment 1
Prioritizing Non Skeletal Muscle Symptoms to Treat

<table>
<thead>
<tr>
<th>Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weaker ability to cough</td>
</tr>
<tr>
<td>Constipation</td>
</tr>
<tr>
<td>Weaker heart pumping</td>
</tr>
<tr>
<td>Frequent waking at night</td>
</tr>
<tr>
<td>Bone fractures</td>
</tr>
<tr>
<td>Lung infections</td>
</tr>
<tr>
<td>Headaches</td>
</tr>
<tr>
<td>Feeling tired</td>
</tr>
<tr>
<td>Non-healthy weight</td>
</tr>
<tr>
<td>Poor attention span</td>
</tr>
<tr>
<td>Depression</td>
</tr>
</tbody>
</table>

Questions on experience with symptoms
Examples:
- Use of cough assist?
- Have you broken a bone?
- Trouble sleeping?
- Lung function
- Heart function

N=155
62% caregivers
38% patients
Results – Symptom Treatment Priorities Stratified

What we learned

Figure 2. Stratified caregiver and patient priorities for symptoms to be treated

- Secondary QOL targets
- Primary QOL targets
- Lifespan targets

- Highest priority
- Poor attention span
- Constipation
- Frequent waking at night
- Feeling tired
- Headaches
- Non-healthy weight
- Depression
- Bone fractures
- Weaker ability to cough
- Lung infections
- Weaker heart pumping
# Experiment 2

**Treatment profile with pulmonary outcomes (meaningful benefit)**

**BW 1 and 2**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
</table>
| Cough Strength             | • Maintained for 10 years  
                            | • Maintained for 2 years  
                            | • Maintained at your current rate |
| Blood Draws                | • No additional blood draws  
                            | • 2 additional blood draws a year  
                            | • 4 additional blood draws a year |
| Knowledge about the drug   | • 0% risk of Diarrhea  
                            | • 20% risk of Diarrhea  
                            | • 50% risk of Diarrhea |
| Lung Infections            | • Very few lung infections  
                            | • Half as many lung infections  
                            | • No reduction in lung infections |

Would you choose this treatment?  
☐ Yes  ☐ No
Results

What we learned from data

• There was little difference found between the preferences of patients and caregivers
• Highest priorities for non skeletal muscle symptoms to treat were cardiac and pulmonary
• Patients and caregivers were willing to accept the risks and burden presented in order to achieve pulmonary benefits
• Maintaining cough strength and less lung infections represent meaningful benefits to both patients and caregivers
• Data was submitted to FDA with Santhera’s regulatory package

Methods:
• Inherent tension between outcomes used in trials and those that matter to patients & families
• “Extensive community engagement re: meaning of a pulmonary benefit and how to describe
• Pilot Testing: Caregivers of young patients found experiment difficult during pilot testing due to thinking about future disease progression
Advancing Methods: Discrete Choice Experiments
Advancing methods – DCE 2016 Research as an Event

Aims:
- Pilot using the DCE
- Testing “Research as an Event”
- Surveying Patients, caregivers, professionals

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscle benefit</td>
<td>Small improvement</td>
</tr>
<tr>
<td></td>
<td>Medium improvement</td>
</tr>
<tr>
<td></td>
<td>Large improvement</td>
</tr>
<tr>
<td>Chance that the drug will work</td>
<td>25% chance</td>
</tr>
<tr>
<td></td>
<td>50% chance</td>
</tr>
<tr>
<td></td>
<td>75% chance</td>
</tr>
<tr>
<td>Extra risk of kidney damage</td>
<td>No extra risk</td>
</tr>
<tr>
<td></td>
<td>10% extra</td>
</tr>
<tr>
<td></td>
<td>20% extra</td>
</tr>
<tr>
<td>Extra risk of fracture</td>
<td>No extra risk</td>
</tr>
<tr>
<td></td>
<td>10% extra</td>
</tr>
<tr>
<td></td>
<td>20% extra</td>
</tr>
</tbody>
</table>

Based on your own opinion, which is the better drug?

<table>
<thead>
<tr>
<th>Drug</th>
<th>Is there muscle benefit?</th>
<th>How many people would benefit?</th>
<th>What is the risk of kidney damage?</th>
<th>What is the extra fracture risk?</th>
<th>In your opinion, which is the better drug?</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Small improvement</td>
<td>50%</td>
<td>No additional risk</td>
<td>20% higher risk</td>
<td>☐</td>
</tr>
<tr>
<td>B</td>
<td>Large improvement</td>
<td>75%</td>
<td>20% higher risk</td>
<td>10% higher risk</td>
<td>☐</td>
</tr>
</tbody>
</table>
Results: Maximum acceptable risk compared to muscle benefit (one level)

Key imperatives that came from pilot:

- Patients appear to be less risk tolerant
- Further explore DCE
- Expand beyond US
Global study of Duchenne treatment preferences

Aims:

• Conduct a multi-country study to quantify treatment preferences in DMD
• Explore global research and advocacy unmet needs
• Administer the EQ-5D and evaluate use in Duchenne
• Develop a survey instrument that can be replicated internationally
• Facilitate the dissemination and implementation of the study finding
Evolving the DCE following community engagement

If both drugs were real and available, which would you choose?

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact on disease progression</td>
<td>Slows for 1 year</td>
</tr>
<tr>
<td></td>
<td>Slows for 3 years</td>
</tr>
<tr>
<td></td>
<td>Slows for 5 years</td>
</tr>
<tr>
<td>Chance that the drug will work</td>
<td>25% chance</td>
</tr>
<tr>
<td></td>
<td>50% chance</td>
</tr>
<tr>
<td></td>
<td>75% chance</td>
</tr>
<tr>
<td>Extra risk of kidney damage</td>
<td>No extra risk</td>
</tr>
<tr>
<td></td>
<td>10% extra</td>
</tr>
<tr>
<td></td>
<td>20% extra</td>
</tr>
<tr>
<td>Extra risk of fracture</td>
<td>No extra risk</td>
</tr>
<tr>
<td></td>
<td>10% extra</td>
</tr>
<tr>
<td></td>
<td>20% extra</td>
</tr>
</tbody>
</table>
## Recruitment

<table>
<thead>
<tr>
<th>Country</th>
<th>Adults</th>
<th>CG of adults</th>
<th>CG of minors</th>
<th>Country total</th>
<th># surveys distributed</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>10</td>
<td>6</td>
<td>13</td>
<td>29</td>
<td>47</td>
<td>62%</td>
</tr>
<tr>
<td>Canada</td>
<td>7</td>
<td>15</td>
<td>30</td>
<td>52</td>
<td>68</td>
<td>76%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>11</td>
<td>16</td>
<td>23</td>
<td>50</td>
<td>53</td>
<td>94%</td>
</tr>
<tr>
<td>UK</td>
<td>24</td>
<td>23</td>
<td>38</td>
<td>85</td>
<td>106</td>
<td>80%</td>
</tr>
<tr>
<td>US</td>
<td>21</td>
<td>25</td>
<td>45</td>
<td>91</td>
<td>111</td>
<td>82%</td>
</tr>
<tr>
<td>France</td>
<td>16</td>
<td>13</td>
<td>5</td>
<td>34</td>
<td>35</td>
<td>97%</td>
</tr>
<tr>
<td>Italy</td>
<td>15</td>
<td>14</td>
<td>33</td>
<td>62</td>
<td>79</td>
<td>78%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>104</td>
<td>112</td>
<td>187</td>
<td>403</td>
<td>499</td>
<td><strong>79%</strong></td>
</tr>
</tbody>
</table>
Preference weights

- For 1 year
- For 3 years
- For 5 years
- 25%
- 50%
- 75%
- No additional
- 10% additional
- 20% additional
- No additional
- 10% additional
- 20% additional

- Slows progression
- Failure rate
- Kidney damage
- Fracture risk

Legend:
- United States
- Italy
- Belgium
- United Kingdom
- Canada
- Netherlands
- France
Maximum acceptable risk

In exchange for 1 year slowing of disease progression

Drug failure rate

Risk of kidney damage

Risk of fracture

Maximum acceptable risk

- United States
- Italy
- Belgium
- United Kingdom
- Canada
- Netherlands
- France
Maximum acceptable risk

In exchange for 1 year slowing of disease progression

- Drug failure rate: 11%
- Risk of kidney damage: 5%
- Risk of fracture: 20%

Bars are 95% confidence intervals
By disease stage: Preference weights

* Between group P-value < 0.05
Maximum acceptable risk

In exchange for 1 year slowing of disease progression

- Drug failure rate
- Risk of kidney damage
- Risk of fracture

- Early ambulatory
- Late ambulatory
- Early non-ambulatory
- Late non-ambulatory

* Between group P-value < 0.05
### Evaluation of preference task

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree or disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The questions were easy to understand</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>The questions were easy to answer</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>My answers showed my real preferences</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>The questions were relevant to me</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

![Chart showing evaluation results](chart.png)

- United States
- Italy
- Belgium
- United Kingdom
- Canada
- Netherlands
- France
DCE evaluation

Bars are ± standard deviation
Acknowledgements

Dr. Holly Peay
Dr. John Bridges
Norah Crossnohere
Pat Furlong
Brian Denger
PPMD Board of Directors
Patients’ Stated Risk Tolerance and Regulatory Decision-Making: Can we trust the numbers?

ISPOR-FDA Summit 2020
September 19, 2020

Juan Marcos Gonzalez
Assistant Professor, Duke School of Medicine
“I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description ["pornography"], and perhaps I could never succeed in intelligibly doing so. But *I know it when I see it*, and the motion picture involved in this case is not that.”
• Is there a Roth test for it?

• Should we *know it when we see it*?

• What is NOT fit for purpose?

• Discussion often relates to types of elicitation methods

• Practical implications
  – Cost
  – Time
Can We Trust the Numbers?

• *Implementation* defines the fit

• What are the implementation issues common to all preference methods that should be considered to establish that something is fit for purpose?
Three Common Steps

1. Ask the right questions
2. Make reasonable assumptions as you interpret the answers
3. Verify the quality of data
Ask the right questions

Clear

Consequential

Meaningful
Clear Questions (Information Stimuli)

- Context
- Tradeoffs
- Data
Example Questions

• Would you accept a treatment for psoriasis if it required accepting a serious outcome?

• Would you accept a treatment for psoriasis that would reduce your lesions if it required accepting a chance of cancer?

• Would you accept a treatment for psoriasis that would reduce your lesions in half but required accepting a 2% chance of cancer in the next 10 years?

• Would you accept a treatment for psoriasis that would reduce your lesions in half but required accepting a 2% chance of cancer in the next 10 years if no other treatment was available?
Framing Effects

- Of 100 people having surgery 90 lived through the postoperative period, 68 were alive at the end of the first year, and 34 were alive at the end of five years.
- Of 100 people having radiation therapy all lived through the treatment, 77 were alive at the end of one year and 22 were alive at the end of five years.

- Of 100 people having surgery 10 died during surgery or the post-operative period, 32 died by the end of the first year, and 66 died by the end of five years.
- Of 100 people having radiation therapy, none died during treatment, 23 died by the end of one year, and 78 died by the end of five years.

Consequential Questions

- Consequential questions are questions that provide respondents an incentive to truthfully reveal preferences (Carson, 1997)
  - Avoid respondent indifference
  - Avoid strategic behavior
- Does not necessarily imply real-world consequences
- Most often deals with the format of the questions and the assumptions respondents are asked to make as they answer
Consequential Questions

• Answers are seen by the agent as potentially influencing an agency’s actions.

• If the agent cares about the outcomes of those actions, the agent should treat the survey questions as an opportunity to influence those actions.

(Carson and Groves, p. 183)
We need your thoughtful answers to help us understand how you feel about gene therapies for SCD.

To make our study a success, we need your help with a problem we have in studies like this. Because participants do not actually have to live with the results of the treatment they select, they often do not think carefully about what they would do if they really had to choose.

If you do not pay attention to the information shown in each question as you would in real life, our results will be wrong. We will not get a true measure of how important various treatment benefits and risks actually are to you and to people like you.
An Example

<table>
<thead>
<tr>
<th>Kidney characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDPI score</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Expected cold ischemia time (at time of surgery)</td>
</tr>
<tr>
<td>Pump parameters (Resistance (R) and Flow (F) levels)</td>
</tr>
<tr>
<td>Nadir and terminal serum creatinine</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Glomerulosclerosis</td>
</tr>
<tr>
<td>Donation after cardiac death</td>
</tr>
<tr>
<td>Diabetes</td>
</tr>
</tbody>
</table>

Would you like to have biopsy information and pump parameters for this kidney?

Would you request a biopsy and pump parameters for this kidney if it means increasing cold ischemia time by 4 hours?
Meaningful Questions

Context  Evidenc e
(In)Dependence of Outcomes

Acceptability of adverse event 1 alone

Acceptability of adverse event 2 alone

Acceptability of adverse events with joint exposure

From Fairchild et al., IAHPR Meeting (Montreal) 2018
Interpreting the answers

• What assumptions are we willing to make to turn patient-preference data into patient-preference information?

• How do we manage the error in our measures?
  – Do we assume no error?
  – Do we make general assumptions about the form of the error with which we collect measures? (e.g., utility maximization process)
    • Model specification
    • Homogeneity of preferences
Expected Utility Theory

Interpolations

From Gonzalez and Boeri (Under review)
Homogeneity

• Adequate evaluation of preference heterogeneity can be challenging due to its empirical nature
  – Might result from a very complicated process
  – Might relate to variables that are difficult to collect

• ISPOR Health Preference Research Special Interest Group is currently conducting a special project on the topic

• Can require a lot of responses from each participant
Confirming the quality of responses

Face Validity
- Elapsed time
- Straight-lining or patterning
- Comprehension (quiz questions)

Content Validity
- Stability (test/retest)
- Within-Set Monotonicity
- Cross-Set Monotonicity
- Transitivity
- Compensatory Preferences (Tradeoffs)
- Scope

Predictive Validity
- Out-of-sample prediction
- Consistency with revealed preferences/behaviors
### Evaluation Summary (N = 30)

<table>
<thead>
<tr>
<th>Evaluation Type</th>
<th>Number of Studies with the evaluation</th>
<th>Range of % Missed</th>
<th>Mean % Missed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeated Question</td>
<td>8</td>
<td>9% -- 76%</td>
<td>35%</td>
</tr>
<tr>
<td>Dominated Pair</td>
<td>11</td>
<td>0% -- 100%</td>
<td>29%</td>
</tr>
<tr>
<td>Cross-set Dominated Pair</td>
<td>14</td>
<td>0% -- 29%</td>
<td>10%</td>
</tr>
<tr>
<td>Transitivity</td>
<td>2</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Dominant Attribute</td>
<td>29</td>
<td>0% -- 92%</td>
<td>30%</td>
</tr>
</tbody>
</table>

At least 75% of the questions

- 6 no dominance
- 11 <10%
- 10 >50%

- Mean = 30.4%
- St Dev = 29.5%

Gonzalez JM. "Evaluating Risk Tolerance from a Systematic Review of Preferences: The Case of Patients with Psoriasis" The Patient, 2018
So, What is Fit for Purpose?

- Should we *know it when we see it*? Should we have a Roth test?
  - Answer: Maybe
  - We do know about the impact of certain decisions during the development of PPI instruments
  - We understand the impact of some assumptions when we make sense of PPI data

- Properly judging fit for purpose will require
  - Paying more attention to our instruments
  - Being explicit about our assumptions
  - Being more critical of our data
Questions?
Applications of Patient Preferences Studies, or
“We measured preferences, now what?”

Bennett Levitan, MD-PhD
Senior Director, Benefit-Risk / Global R&D Epidemiology
Janssen R&D, LLC

Session 3: Methodologic Issues for PPI Studies
Virtual ISPOR-FDA Summit 2020
September 29, 2020
Disclosure / Conflict of Interests Statement

• Dr. Levitan is an employee of Janssen Research and Development, LLC. He is a stockholder in Johnson & Johnson. He also owns stock in a portfolio that at times includes other pharmaceutical and health care-related companies.

• The views and opinions expressed in this presentation are those of the individual presenter and do not necessarily reflect the views of Janssen Research & Development, the United States Food and Drug Administration or the International Society for Pharmacoeconomics and Outcomes Research, Inc.
Where Can Patient Preferences Decisions in the Lifecycle? (What research questions can be addressed?)

- **Commercial viability / Patient needs**: What endpoints do patients care most about? What level/rate of endpoints are critical to patients? What is the relative importance of benefits, risks and other treatment features to patients? How do patients vary in these properties (heterogeneity)? Are there distinct subgroups?

- **Trial design**: What endpoints do patients care most about? What level/rate of endpoints are critical to patients?

- **TPP**: What is the relative importance of benefits, risks and other treatment features to patients? Maximum acceptable risk, minimum required benefit, choice share?

- **Approval & reimbursement**: How do patients vary in these properties (heterogeneity)? Are there distinct subgroups? Are there important differences between stakeholders? Shared decision-making

- **Ph 2a/b**: What endpoints do patients care most about? What level/rate of endpoints are critical to patients?

- **Ph 3**: What is the relative importance of benefits, risks and other treatment features to patients? Maximum acceptable risk, minimum required benefit, choice share?

- **Reg**: How do patients vary in these properties (heterogeneity)? Are there distinct subgroups? Are there important differences between stakeholders? Shared decision-making

- **Post-approval**: How do patients vary in these properties (heterogeneity)? Are there distinct subgroups? Are there important differences between stakeholders? Shared decision-making

TPP = Target Product Profile
Many Approaches And Complications → Application of Preferences to Clinical Data and Decision-making Is Not Always Clear or Straightforward

**Approaches**

- Maximum acceptable risk
- Choice share
- Net clinical benefit
- Relative importance
- Multi-criteria decision analysis
- Visualization
- Effect size assessment
- Minimum acceptable benefit
Many Approaches And Complications → Application of Preferences to Clinical Data and Decision-making Is Not Always Clear or Straightforward

**Approaches**
- Maximum acceptable risk
- Choice share
- Net clinical benefit
- Relative importance
- Multi-criteria decision analysis
- Visualization
- Effect size assessment
- Minimum acceptable benefit

**Complications**
- Heterogeneity of clinical results
- Heterogeneity of preferences
- Uncertainty in clinical results
- Uncertainty in preferences
- Dependency between endpoints
- Dependency between preferences
- Many benefits and many harms
- Non-linear preferences
### Three Broad Classes for Applying Preference Data to Inform Decision-Making

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences alone</td>
<td>Relative importance, maximum acceptable risk, minimum acceptable benefit, effect size</td>
</tr>
<tr>
<td>Preferences + clinical data viewed jointly</td>
<td>Graphical techniques (parallel display of clinical data and preferences)</td>
</tr>
<tr>
<td>Preferences + clinical data mathematically combined</td>
<td>Choice share, net clinical benefit, multi-criteria decision analysis (MCDA) modeling, …</td>
</tr>
</tbody>
</table>
Relative Importance – Qualitative Preferences
Determining Which Endpoints are Most Critical

Value Tree for Acute Migraine

Benefits
- ↓ Pain
  - Rapid onset
  - Headache relief
  - Pain-free response
  - Sustained response
- ↓ Sensitivity
  - Reduced sensitivity to sound & light
- ↓ Other
  - Reduction in functional disability
  - Reduction in nausea or vomiting

Benefit-Risk Balance

Risks
- ↑ Individual Risks
  - CNS adverse events
  - “Chest-related” adverse events
  - Myocardial infarction

Relative Importance – Qualitative Preferences
Determining Which Endpoints are Most Critical

Value Tree for Acute Migraine

Benefits

- Pain
  - Pain-free response
  - Sustained response
- Sensitivity
  - Reduced sensitivity to sound & light

Benefit-Risk Balance

- Other
  - Reduced nausea or vomiting
  - Reduced sensitivity to sound & light

Risks

- Individual Risks
- Other Risks

Physician: “I was really struck that you threw out the parameter that we focused the most on. We thought that if you were going to have the risk of a heart attack, you should really get rid of your migraine, period.”

Preferences
Determining Which Endpoints are Most Critical

- Fragile-X Syndrome
  - Rare genetic condition impacting development
  - Learning and intellectual disabilities, cognitive impairment, behavioral challenges (ADHD, autism, social anxiety), physician features
  - No cure – educational, therapeutic support

- Preference study conducted to prepare for phase 3 study
  - Intent was to identify which endpoints were most important to patients
  - Survey administered to family members, given cognitive limitations
Preference Survey Identified Large Gap Between Clinician and Patient Caretaker Beliefs on Endpoint Importance

N = 614

Preference Survey Identified Large Gap Between Clinician and Patient Caretaker Beliefs on Endpoint Importance

N = 614

Preference Survey Identified Large Gap Between Clinician and Patient Caretaker Beliefs on Endpoint Importance

Clinical and commercial perspective of the most important endpoints

Caretaker perspective of the most important endpoints

N = 614

Benefit-Risk Assessment: Maximum Acceptable Risk (of death or permanent severe disability due to stroke)

Useful approach for simpler B-R assessments and target product profiles

What if there are several risks?
What if the level of benefit is uncertain?
How can we account for population heterogeneity in preference?

Can tune the joint maximum acceptable risk threshold to allow for different proportions of the population to consider $B > R$. 
Can use the same approach with more complex maximum acceptable risk thresholds.
Effect Size (mock example)

Effect Size (mock example)

Effect Size (mock example)

Novel “time-to” Efficacy Endpoint

Heart Attack Risk

Preference change for the accepted effect size

Well-accepted effect size (0.5%)
Effect Size (mock example)

Effect Size (mock example)

### Three Broad Classes for Applying Preference Data to Inform Decision-Making

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences alone</td>
<td>Relative importance, maximum acceptable risk, minimum acceptable benefit, effect size</td>
</tr>
<tr>
<td>Preferences + clinical data viewed jointly</td>
<td>Graphical techniques (parallel display of clinical data and preferences)</td>
</tr>
<tr>
<td>Preferences + clinical data mathematically combined</td>
<td>Choice share, net clinical benefit, multi-criteria decision analysis (MCDA) modeling, …</td>
</tr>
</tbody>
</table>
### Forest Plot with Endpoints Sorted by Preference

**Atrial Fibrillation Example (mock data)**

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Study Drug</th>
<th>Comparator</th>
<th># events / 10,000 patient-years</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>187</td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>Disabling stroke</td>
<td>39</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>Non-CNS systemic embolism</td>
<td>3</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Non-disabling stroke</td>
<td>80</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>90</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Major bleeding</td>
<td>361</td>
<td>346</td>
<td></td>
</tr>
<tr>
<td>Non-major clinically relevant bleeding</td>
<td>1179</td>
<td>1135</td>
<td></td>
</tr>
</tbody>
</table>

**Rate Difference per 10,000 patient-years**

- **Most severe**
  - Major bleeding: 361 - 346 = 15
  - Non-major clinically relevant bleeding: 1179 - 1135 = 44
  - Myocardial infarction: 90 - 111 = -21
  - Non-disabling stroke: 80 - 78 = 2
  - Non-CNS systemic embolism: 3 - 18 = -15
  - Disabling stroke: 39 - 51 = -12
  - All-cause mortality: 187 - 221 = -34

- **Least severe**
  - Major bleeding: 361 - 346 = 15
  - Non-major clinically relevant bleeding: 1179 - 1135 = 44
  - Myocardial infarction: 90 - 111 = -21
  - Non-disabling stroke: 80 - 78 = 2
  - Non-CNS systemic embolism: 3 - 18 = -15
  - Disabling stroke: 39 - 51 = -12
  - All-cause mortality: 187 - 221 = -34

Favors Study Drug | Favors Comparator
Forest Plot with Endpoints Positioned by Preference

Atrial Fibrillation Example (mock data)

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Study Drug</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>187</td>
<td>221</td>
</tr>
<tr>
<td>Disabling stroke</td>
<td>39</td>
<td>51</td>
</tr>
<tr>
<td>Non-CNS systemic embolism</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Non-disabling stroke</td>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>90</td>
<td>111</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>361</td>
<td>346</td>
</tr>
<tr>
<td>Non-major clinically relevant bleeding</td>
<td>1179</td>
<td>1135</td>
</tr>
</tbody>
</table>

# events / 10,000 patient-years

Most severe

Least severe

Rate Difference per 10,000 patient-years

Favors Study Drug  Favors Comparator
### Atrial Fibrillation Example (mock data)

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Study Drug</th>
<th>Comparator</th>
<th># events / 10,000 patient-years</th>
<th>Rate Difference (95% CI) / 10,000 pat-years</th>
<th>Rate Difference</th>
<th>Preference Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>187.8</td>
<td>222.0</td>
<td>-34.2 (-71.62, 3.29)</td>
<td>Blue</td>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td>Disabling stroke</td>
<td>38.5</td>
<td>50.2</td>
<td>-11.8 (-29.31, 5.81)</td>
<td>Blue</td>
<td></td>
<td>0.811</td>
</tr>
<tr>
<td>Non-CNS systemic embolism</td>
<td>4.5</td>
<td>19.4</td>
<td>-14.9 (-24.29, -5.53)</td>
<td>Blue</td>
<td></td>
<td>0.420</td>
</tr>
<tr>
<td>Non-disabling stroke</td>
<td>78.8</td>
<td>76.7</td>
<td>2.0 (-21.16, 25.19)</td>
<td>Blue</td>
<td></td>
<td>0.359</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>90.8</td>
<td>111.7</td>
<td>-20.9 (-47.33, 5.59)</td>
<td>Blue</td>
<td></td>
<td>0.275</td>
</tr>
<tr>
<td>Major bleeding</td>
<td>359.6</td>
<td>345.4</td>
<td>14.2 (-35.30, 63.73)</td>
<td>Red</td>
<td></td>
<td>0.267</td>
</tr>
<tr>
<td>Non-major clin relevant bleeding</td>
<td>1180.4</td>
<td>1136.6</td>
<td>43.8 (-50.23, 137.77)</td>
<td>Red</td>
<td></td>
<td>0.159</td>
</tr>
</tbody>
</table>
Three Broad Classes for Applying Preference Data to Inform Decision-Making

<table>
<thead>
<tr>
<th>Class</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences alone</td>
<td>Relative importance, maximum acceptable risk, minimum acceptable benefit, effect size</td>
</tr>
<tr>
<td>Preferences + clinical data viewed jointly</td>
<td>Graphical techniques (parallel display of clinical data and preferences)</td>
</tr>
<tr>
<td>Preferences + clinical data mathematically combined</td>
<td>Choice share, net clinical benefit, multi-criteria decision analysis (MCDA) modeling, …</td>
</tr>
</tbody>
</table>
Choice Share: What Proportion of Patients Would Use One Treatment vs Others

- Preference study can give the proportion of patients who would use one treatment vs others
- Supported CDRH approval decision on weight loss device
- Can be applied to population as a whole or subgroups with greater/lesser risk tolerance

<table>
<thead>
<tr>
<th>Weight loss device profile</th>
<th>% Respondents who judged better than a no-device alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual device A</td>
<td>5.0 %</td>
</tr>
<tr>
<td>Benefit: 5 % TBWL</td>
<td></td>
</tr>
<tr>
<td>Risk: 1 % chance of death</td>
<td></td>
</tr>
<tr>
<td>Type of surgery: laparoscopic surgery</td>
<td></td>
</tr>
<tr>
<td>Dietary restriction: eat 1/4 cup of food at a time</td>
<td></td>
</tr>
<tr>
<td>Weight-loss duration: 36 months</td>
<td></td>
</tr>
<tr>
<td>Minor side-effect duration: 36 months</td>
<td></td>
</tr>
<tr>
<td>Chance of hospitalization with surgery: 5 %</td>
<td></td>
</tr>
<tr>
<td>Comorbidity: no improvement</td>
<td></td>
</tr>
<tr>
<td>Gastric band</td>
<td>11.6 %</td>
</tr>
<tr>
<td>Benefit: 13 % TBWL</td>
<td></td>
</tr>
<tr>
<td>Risk: 1 % chance of death</td>
<td></td>
</tr>
<tr>
<td>Type of surgery: laparoscopic surgery</td>
<td></td>
</tr>
<tr>
<td>Dietary restriction: eat 1/4 cup of food at a time</td>
<td></td>
</tr>
<tr>
<td>Weight-loss duration: 5 years</td>
<td></td>
</tr>
<tr>
<td>Minor side-effect duration: 5 years</td>
<td></td>
</tr>
<tr>
<td>Chance of hospitalization for side effects requiring surgery: 5 %</td>
<td></td>
</tr>
<tr>
<td>Comorbidity: no improvement</td>
<td></td>
</tr>
<tr>
<td>Virtual device B</td>
<td>27.5 %</td>
</tr>
<tr>
<td>Benefit: 20 % TBWL</td>
<td></td>
</tr>
</tbody>
</table>

Multi-criteria Decision Analysis

- Endpoints converted to a normalized scale
- Weights applied to each endpoint
- Cumulative impact summed and displayed graphically

Modeling: Accounting for Uncertainty in Clinical Data and Preference Weights *

Favors Treatment A  Favors Treatment B

$p(benefit > risk \text{ for treatment A vs B}) = 97\%$

Mean net clinical benefit = -48 (-95,0.3)

* Stochastic multi-attribute acceptability analysis
So are we all set?

- How can patient preferences support a B-R decision?

- Can we calculate maximum acceptable risk for a given benefit and compare to the probability of the harm in the clinical data

- Not always …
So are we all set?

- How can patient preferences support a B-R decision?
- Can we calculate maximum acceptable risk for a given benefit and compare to the probability of the harm in the clinical data?
- Not always …
Preference Study for Triptans in Acute Migraine

- Triptans are vasoconstrictors → Risk of myocardial infarction (MI)
- Attributes of greatest relative importance in DCE were activity limitations and MI
- Consider maximum acceptable risk (MAR) for treatment related MI
- MAR = additional 0.2% MI / year in exchange for severe → no limitations

Gonzalez, et. al., ISPOR 17th Annual Intl Mtg, Washington, DC, June 2012.
Applying the Maximum Acceptable MI Risk:

A bit more complex than it seems

Severe to none on activity limitations → would accept 0.2% annual chance of MI
But not everyone benefits
Example:

Triptan: 40% migraines with no activity limitations
NSAID: 56% migraines with no activity limitations

16% with this benefit, will accept a 0.2% annual chance of MI
84% without this benefit, will not accept the MI from the benefit

→ clinical trial sample MAR overall is (0.2%) x 0.16 = 0.032%

So … if MI risk difference < 0.032%, benefits-exceed risks

But what about those who benefit to a lesser degree? → more complex approach with multiple MAR assessments
<table>
<thead>
<tr>
<th>Class</th>
<th>Flexibility</th>
<th>Technical Complexity</th>
<th>Ease of Communication</th>
<th>Clarity / Transparency</th>
<th>Health Authority Acceptance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences alone</td>
<td>Low - Medium</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Medium - High</td>
</tr>
<tr>
<td>Preferences + clinical data viewed jointly</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Preferences + clinical data mathematically combined</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low - Medium</td>
</tr>
</tbody>
</table>
Take Away Messages

• Numerous approaches to using patient preference data to inform decisions
• Use the simplest approach that addresses the research questions
  – Ideally those approaches that use preference data independently
  – For more complex questions, use parallel display of clinical and preference data
  – Mathematical combination approaches are generally used in the academic literature, though HAs are considering or starting to use them
• Often use a combination of approaches
• Real world applications are often not straightforward
To ask a question, either:

1. Use the live Q&A feature in the app
2. Click on the thought bubble icon in the webcast window
Virtual ISPOR-FDA Summit 2020

Break

Program will return at 3:15PM EDT